

THE CONDITIONED REINFORCEMENT OF REPEATED ACQUISITION¹

STEVEN R. HURSH²

WALTER REED ARMY INSTITUTE OF RESEARCH

Three monkeys were trained to emit a chain of three responses on three separate levers in a set of six levers to obtain food. The chain producing food (correct chain) was changed each day. During a trial, a press on any lever produced a feedback stimulus; a press on a correct lever produced an additional distinctive stimulus; the third correct press produced a food pellet. Test sessions in which either the food or the distinctive stimuli were removed were interspersed with baseline sessions. In tests without food presentations, the subjects acquired the correct chain rapidly, with a level of accuracy comparable to baseline. Removing the distinctive stimuli for either the first or second member of the correct chain greatly retarded acquisition of that member of the chain. Removing all distinctive stimuli often reduced accuracy throughout the chain to chance level, even though food was presented following each correct chain. These results were interpreted as evidence that the distinctive stimuli presented after correct responses functioned as conditioned reinforcers. Reductions in accuracy following an omitted distinctive stimulus indicated that they were also discriminative stimuli for correct responding in their presence.

Key words: conditioned reinforcement, repeated acquisition, heterogeneous chain schedule, stimulus control, monkeys

The methodology for this study was an adaptation of the repeated acquisition procedure first described by Boren (1963). The subject must emit a specified chain or sequence of responses on different manipulanda to obtain food reinforcers (a heterogeneous chain, Nevin, 1973). The correct chain of responses can be altered frequently, thereby requiring the subject repeatedly to acquire different chains of responses. After preliminary training, subjects typically demonstrate a stable state of rapid acquisition each time a new chain is required (Boren and Devine, 1968; Thompson, 1971).

The rate of acquisition can be influenced by stimulus events occurring during the chain. For example, Boren and Devine (1968) and Boren (1969) found that errors during acquisition

could be reduced by 90% by presenting a 1-sec timeout after each error. Thompson (1975) demonstrated a similarly large decrease in errors when different discriminative stimuli signalled each link of the chain. These results suggest that the addition of either discriminative stimuli or differential consequences for correct and incorrect responses improves accuracy during acquisition.

The present experiment was designed to separate the enhancing effects of discriminative stimuli from the strengthening effects of differential consequences during the acquisition of a chain. A distinctive stimulus was presented after each correct member of the chain. After stable repeated acquisition was attained with these stimuli after each correct response, they were systematically omitted after some or all of the members of the chain to analyze their influence on the accuracy of responding. These stimuli could be said to function as reinforcers if they increased the probability of the preceding new response and they could be said to function as discriminative stimuli if they altered the probability of subsequent new responses. These two functions could be separately analyzed because each member of the chain was topographically distinct. This study explored this dual function of each distinctive

¹In conducting this research, the investigators adhered to the "Guide for the Care and Use of Laboratory Animals DHEW Publication No. (NIH) 73-23, as Prepared by the Institute of Laboratory Animal Resources, National Research Council.

²Reprints may be obtained from the author, Department of Experimental Psychology, Walter Reed Army Institute of Research, Washington, D.C. 20012. The author thanks Andrew Allen and Michael Ruggiero for their efficient assistance conducting this research and Wilhelmina Taylor for her special assistance preparing the manuscript.

stimulus presented after each member of a three-link chain.

METHOD

Subjects

Two rhesus monkeys (*Macaca mulatta*) weighing approximately 5.2 kg and 6.5 kg (SM 4 and SM 5 respectively) and one cynomolgus monkey (*Macaca fascicularis*) weighing approximately 3.0 kg (SM 6) served throughout the experiment. Each monkey obtained its total daily ration of food pellets (Noyes banana-flavored 750-mg pellets, formula L) as reinforcers during the sessions: SM 4, 100 pellets; SM 5, 150 pellets; and SM 6, 80 pellets. These rations were approximately 80% of their free intake on a continuous fixed-ratio three (FR 3) schedule. Each subject also received a quarter of an orange or apple each morning.

Apparatus

The subjects lived continuously in cages mounted next to work panels. Two types of work panel were used, one with levers (SM 4) and another with push plates (SM 5 and SM 6). Both were arranged similarly and will not be described separately (see Figure 1). The panels had two rows of three keys (levers or push plates). Centered below the six keys was a circular opening to a food cup. Above the top row of keys on the left was a response button beside a water tube. Located above each lever was a 3-cm circular window with a stimulus projector (Industrial Electric inline projector model 10) mounted behind it. Several colors and patterns could be projected onto the translucent screen singly or in combination. The push plates could be transilluminated yellow or green and a white spot in the upper portion of the key could be superimposed on a green key. The food cup was illuminated by two clear bulbs. All lamps were 28 V dc. The three work panels were connected to solid-state programming and recording equipment in an adjacent room. General illumination was provided 12 hr per day from 0500 to 1700 hr.

Procedure

Preliminary training. For simplicity of explanation, no distinction is made between levers and lights (SM 4) and transilluminated keys (SM 5 and SM 6). Initially, all subjects

were trained to press one lighted key to receive pellets; the other five keys were dark and provided no scheduled consequences. After two sessions, the response requirement was raised to three presses (FR 3). After several additional sessions, a trials procedure was introduced starting at 1400 hr each day, seven days a week, and terminating after the daily ration of pellets was delivered (see Subjects). Each trial started with all six keys lighted yellow (see Figure 1). Pressing a key would change the keylight to green (SM 5 and 6) or a circle (SM 4), stimuli henceforth referred to as feedback stimuli. Pressing three different keys in succession illuminated three feedback stimuli, produced a food pellet in a lighted food cup, and, after 1 sec, terminated the trial and the three feedback stimuli. After a 5-sec intertrial interval with all keys dark, the six yellow lights again appeared to start a new trial. During any trial, a second press on any one key terminated the trial without a food pellet and initiated a 10-sec intertrial interval. This procedure was in effect for 10 to 27 sessions.

Baseline procedure. After stable three-key chaining was acquired and repetitive pressing on one key was infrequent, the baseline procedure was gradually introduced. One three-key chain was designated as correct each day. Each correct key press (a response on a key in its proper position in the chain) resulted in presentation of the feedback stimulus and a superimposed distinctive stimulus (a white spot of light centered in the top half of the key for SM 5 and SM 6; a white cross for SM 4). An incorrect key press resulted in presentation of the feedback stimulus alone. The stimuli illuminated on the other keys were unaffected. The feedback stimuli and any superimposed distinctive stimuli remained illuminated on the keys until the end of the trial. The trials ended after three keys were pressed. If all three responses were correct and three distinctive stimuli were illuminated, a food pellet was delivered, accompanied by a 0.5-sec illumination of the food cup. If all three responses were not correct and fewer than three distinctive stimuli were illuminated, no pellet was delivered and the food cup remained dark. Whether or not the chain was correct, all keys were darkened 1 sec after the third response. After a 5-sec intertrial interval, the next trial began with keys again illuminated yellow. If any key was pressed more than once during a trial, the

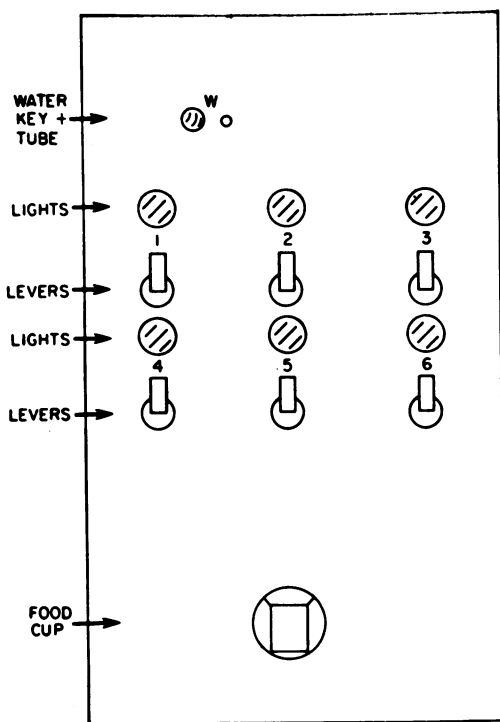


Fig. 1. The work panel (not to scale) for SM 4 showing the six levers and lights, the food cup, and the water dispenser. The intelligence panel for SM 5 and SM 6 was arranged similarly, except that transilluminated keys were substituted for levers and lights.

trial terminated immediately and was followed by a longer, 10-sec intertrial interval. This error seldom occurred and these short trials were not included in the tabulation of results. Thus, all valid trials contained three and only three responses on three keys, some or all of which could be correct. This ensured an equal number of opportunities to respond in each position of the chain throughout all conditions of the experiment.

The correct chain was changed only after acquisition was demonstrated in one daily session, *i.e.*, the subject earned the entire day's ration by performing the correct chain. Early in training, new chains often differed from previous chains by only one response. Occasionally, pressing certain keys was made ineffective (no scheduled consequence) when pressing was sufficiently persistent to prevent acquisition of the new chain. These special procedures were required during 18 to 30 sessions. After initial training on the baseline procedure, a different chain was designated correct each day. The

chains designated correct from day to day encompassed the entire set of 120 possible three-key sequences, but a new chain had to satisfy the following restriction: (1) it could not be composed of presses on the same three keys of the previous day, and (2) no one key could appear in the same position on two adjacent days. Chains were designated by numbering the six keys from left to right and top to bottom (see Figure 1). A chain designated 2-5-6, for example, could not be followed the next day by a correct chain of 6-2-5 (restriction 1) or 3-5-4 (restriction 2).

Test conditions. After repeated acquisition with distinctive stimuli was stable from day to day, *i.e.*, the subject was earning the entire daily ration during the sessions and the number of errors per day showed no upward or downward trend, a test condition was conducted every three days. To determine if the distinctive stimuli superimposed on the feedback stimuli were reinforcers, they were systematically removed during test sessions to examine their effects on the preceding member of the chain. All trials still required three responses; only the stimulus consequences were changed. Most tests involved removing for a session either three, two, or one of the distinctive stimuli. When one distinctive stimulus was removed, for example, the correct first, second, or third member of the chain produced the feedback stimulus without the superimposed distinctive stimulus. Each type of test session was conducted three or four times, intermixed with other tests in no systematic order. Tests conducted with no distinctive stimuli for any correct responses inhibited acquisition more than others and were conducted more often to obtain a satisfactory sample of performance. When a subject did not complete a session, it was terminated the next day at 0800 hr and supplemental food was given. The actual number of tests of each type conducted with each subject is shown in Table 1.

In addition to tests conducted without distinctive stimuli, three tests with SM 5 and SM 6 were conducted with the distinctive stimuli, but without food or food-cup illumination. These tests were identical to baseline conditions, except that the consequences at the end of a trial were identical for correct and incorrect chains; however, each correct response in the chain produced the appropriate distinctive stimulus. These sessions were terminated after

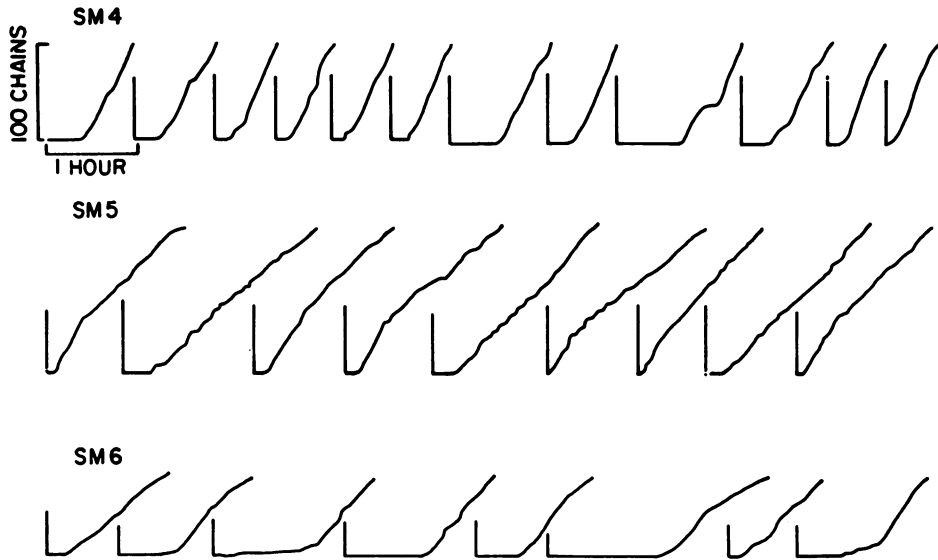


Fig. 2. Cumulative records of correct chains for the three subjects from representative baseline sessions. The pen reset only between sessions, so that each excursion of the pen was a single acquisition curve.

the usual number of correct chains and were immediately followed by a session the same day in which food was again provided for correct chains. This maintained the usual daily food intake without introducing free supplemental food after the test sessions. These tests were conducted as the final conditions of this experiment. SM 4 did not receive these tests because it had advanced to another experiment.

An exact record of each session's chains was recorded on punched paper tape for computer analysis and, in addition, gross totals of errors, corrects, reinforcers, and durations of the three links of the chain were recorded on counters. Cumulative records of correct chains were made.

RESULTS

Acquisition with Distinctive Stimuli and Food (Baseline)

Trials terminated by repeated presses on a single key were infrequent and have not been included in the tabulation of results. Figure 2 shows sample cumulative records from each subject. The pen stepped upward with each correct chain as the paper moved during each trial. The pen reset only at the beginning of each session. Flat periods in the records indicated either a number of trials without a correct chain or trials with a low rate of respond-

ing. For the most part, observations of the subjects indicated that flat periods early in the session were occasions of repeated errors; acquisition was usually rapid and occurred suddenly. The subjects' performance usually made an abrupt transition from no correct chains to consistently correct chains similar in form to acquisition of lever pressing by rats reported by Skinner (1938). These records appear to have a more rapid transition point than those of pigeons acquiring four-response chains (Thompson, 1975).

A useful summary measure of performance on this task is the per cent correct responses for each member of the chain. This was computed by dividing the total number of correct key presses for each member of the chain by the total number of chains in the session and multiplying by 100. If the three members of the correct chain were acquired simultaneously and without error (no incorrect chains), then a score of 100% would be achieved for each member. If no acquisition occurred and responding was random, then a score of 16.7% would be observed for each member, the chance level of correct responding.³

³The calculation of the random probability of correct responding in each link excluding trials in which a single key was pressed twice is as follows: (a) The probability of a correct response in the first link, p_1 , is $1/6$ or 0.167. (b) The probability of a correct response

Figure 3 shows the per cent correct for each member of the chain for each condition of the experiment, averaged across test and baseline sessions. The lines extending above each bar indicate the best single performance for that member of the chain in each condition, a measure of variance. Baseline performance is shown at the far left of the figure. These baseline scores were based on the average of the sessions immediately preceding each test condition (20 sessions for SM 4, 26 for SM 5, and 29 for SM 6). The three filled bars show the per cent correct for the three members of the chain when a distinctive stimulus followed each correct response. Notice that for each subject the three members of the chain were acquired with nearly equal accuracy, although there was a consistent tendency for the first response to be more accurate than the response preceding food presentation. In 57 of 75 baseline sessions, the first response was more accurate than the third response ($p < 0.01$, given an expected frequency of 38). This difference appears small in terms of overall accuracy in Figure 3 because differences in acquisition accuracy early in the session were added to nearly perfect and uniform accuracy of performance late in the session.

Acquisition with Distinctive Stimuli and No Food

SM 5 and SM 6 were given three tests, each separated by two baseline sessions, in which correct responses produced distinctive stimuli but correct chains were not followed by food or food-cup illumination. These tests were carried out at the end of the experiment but the results are presented here as a reference point for evaluation of the tests without distinctive stimuli.

The second set of bars in Figure 3 shows the per cent correct for each member of the chain as an average of the three test sessions. The levels of accuracy and resulting acquisition

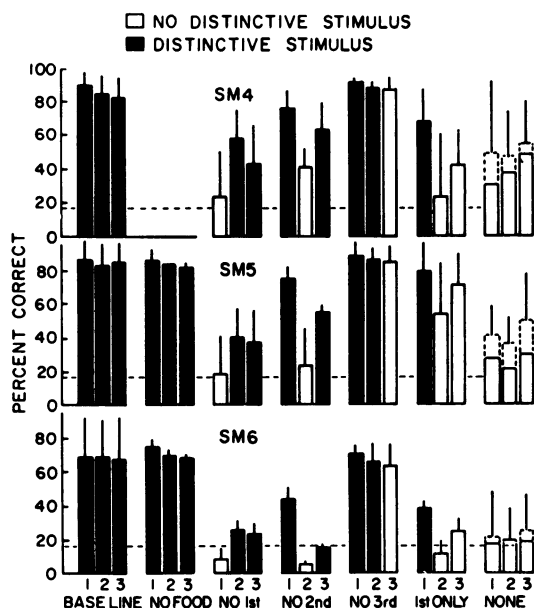


Fig. 3. The per cent correct responding in each link of the chain for each subject across conditions (see text for description of calculations). Filled bars are for links in which correct responses produced distinctive stimuli, open bars are links without distinctive stimuli. Lines extending above each bar indicate the best single performance in that condition in that link. The horizontal dashed line indicates the chance level of accuracy. See text for explanation of the labels for each condition and the dotted bars in the no distinctive stimulus ("None") condition.

during these tests were nearly indistinguishable from baseline. Clearly, the presentation of food was not necessary for the acquisition of new chains when test sessions were interspersed with baseline sessions. One would expect, however, that continued exposure to the distinctive stimuli without food would eventually terminate responding.

Figure 4 shows cumulative curves of the number of correct responses for each member of the chain (dots—first member; circles—second member; triangles—third member). The first and third tests are shown; the results of the second were similar. The solid diagonal of each panel is the locus of 100% correct. The dashed line across the records marks the first correct chain in the session. It is clear that acquisition of each member of the chain and of the chain as a whole was rapid; these records are nearly identical to baseline sessions with food presented. Sessions conducted with no food and without distinctive stimuli (not shown) gave no evidence of acquisition. The

in the second link, P_2 , is the probability that the correct key was not pressed in the first link, $1 - p_1$, times the probability that the correct key is selected, $1/5$, or $(1 - 0.167) \times (1/5) = 0.167$. (c) The probability of a correct response in the third link, p_3 , is the probability that the correct key was not pressed in the first or second link, $1 - (p_1 + p_2)$, times the probability of a correct response, $1/4$, or $[1 - (0.167 + 0.167)] \times (1/4) = 0.167$. (d) Therefore, the random probability of each correct response was 0.167.

tendency to acquire the first member before the third was apparent in these tests without food, as in baseline sessions (five of six cases, see Figure 4, but the small sample of only six sessions increases the statistical probability of this proportion to greater than 0.01).

Table 1 shows the rate of emitting chains (chains per minute) excluding intertrial-interval time. The rate of responding when no food was presented was about equal to the rate under baseline conditions.

Acquisition with Food but Without Distinctive Stimuli

Elimination of all distinctive stimuli had the most retarding effect on all test conditions, far more disruptive than elimination of food reinforcers. On the average, these sessions were much longer than baseline sessions and the rate of emitting chains was lower than in any other condition (see Table 1). As a consequence, the rate of reinforcement was also lowest in this condition. Often, no correct chains occurred during an extended session of 18 hr despite as many as 1200 trials. Since food presentation was the only source of reinforcement, tests were carried out with each subject until at least three sessions occurred with more

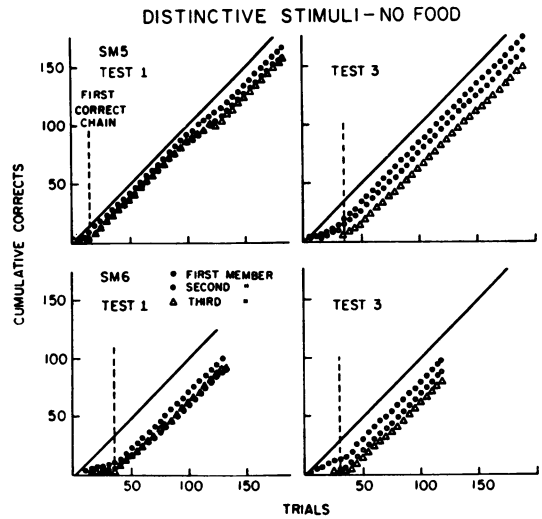


Fig. 4. Cumulative curves of correct responses in each link as a function of blocks of five trials during the condition with distinctive stimuli but no food. The diagonal line is the slope of perfect performance. These records were constructed from a computer summary of the results of each trial in each session; duration of each trial is not reflected in the scaling of the x-axis.

than 20 reinforcements. The analysis of percent correct for this condition is shown as unfilled bars in the right section of Figure 3, in

Table 1

Summary of conditions, number of sessions, and, in parentheses, the number of completed sessions, and the mean results of these sessions: total chains, chains per minute (excluding intertrial-interval time), total correct chains, and total correct responses in the first, second, and third positions of the chain.

Subject	Condition	Sessions (Completed)	Total Chains	Chains per Minute	Correct Chains	Correct First	Correct Second	Correct Third
SM 4	Baseline	20 (20)	132	16.5	100	117	110	104
	No S ^r	7 (4)	819	1.5	58	137	250	340
	No 1st S ^r	3 (2)	450	5.7	69	89	249	197
	No 2nd S ^r	3 (3)	331	12.7	100	251	124	140
	No 3rd S ^r	3 (3)	119	17.0	100	108	105	103
	1st S ^r only	4 (4)	1068	15.9	100	632	146	416
SM 5	Baseline	26 (26)	189	7.9	150	164	156	159
	No food	3 (3)	186	8.1	150	162	158	154
	No S ^r	10 (3)	829	1.2	49	195	140	203
	No 1st S ^r	3 (2)	1248	2.5	107	174	465	446
	No 2nd S ^r	3 (2)	1173	2.7	217	845	168	653
	No 3rd S ^r	3 (3)	183	7.6	150	160	156	154
	1st S ^r only	4 (4)	588	7.8	150	417	177	353
SM 6	Baseline	29 (29)	136	11.3	80	91	90	89
	No food	3 (3)	127	9.8	80	95	88	86
	No S ^r	10 (2)	862	1.2	19	114	154	127
	No 1st S ^r	4 (2)	777	2.6	41	60	199	175
	No 2nd S ^r	4 (0)	1231	2.9	9	539	57	185
	No 3rd S ^r	4 (4)	143	11.9	80	101	93	101
	1st S ^r only	4 (2)	1364	2.5	45	516	140	332

terms of both an average of all these tests (solid outline bars) and as an average of only those tests with at least 20 reinforcers (dotted outline bars). Acquisition with no food but with distinctive stimuli (previous condition) was vastly superior to acquisition with no distinctive stimuli but with food, even if we consider only those sessions in which 20 or more pellets were delivered. Only SM 4 performed in this condition with accuracy generally better than chance (16.7%, see Figure 3). It completed four of the seven tests conducted without distinctive stimuli. Of the 10 such tests with SM 5 and SM 6, they completed only three and two sessions, respectively, and each received over 20 reinforcers in one additional session. Performance by SM 5 and SM 6 was generally random and it was difficult to discern any consistent difference in the rate of acquiring the various members of the chain. If we consider only those tests with a significant number of primary reinforcers (over 20), accuracy was somewhat higher, especially with SM 4 and SM 5, and the subjects showed slightly superior accuracy for the third member of the chain.

The three curves in Figure 5 shows the simultaneous accumulation of correct responses in the three links of the chain across blocks of 50 trials during one completed session by each subject. Presentation of the fifth and tenth food pellets is marked by dashed vertical lines; the slopes of perfect and chance performance are indicated. The slopes of the curves for SM 4 and SM 6 before the fifth food pellet were not equal, presumably a result of historical factors leading to a greater tendency to press certain keys. Acquisition is most accurately inferred from an acceleration in correct responding, not from the absolute level of correct responding. Based on this assumption, acquisition of the three members of the chain occurred at different points in the sessions. For example, in the session depicted for SM 4, the curve for the third member began a steady acceleration between the fifth and tenth reinforcer; later, after the tenth reinforcer, the curve for the second member showed a small but consistent increase in slope; finally, after 100 trials, the curve for the first member of the chain showed a similar acceleration. Each subject showed several different patterns of acquisition, exemplified by the curves for SM 5 and SM 6. In the session depicted for SM 5, the

first and third members of the chain were acquired first at about the same time; the second member of the chain was not acquired until about 250 trials later. Note that when the second member of the chain was finally acquired, the rate of the first and third members also increased, probably as a result of the increased rate of food reinforcement following correct chains (three correct responses). In the session depicted for SM 6, the rate of correct responding was greater than chance for the second and third members of the chain starting early in the session, but little improvement in performance occurred throughout the session. The only clear acquisition occurred for the first member of the chain after the fifth food reinforcer. These records of test sessions with no distinctive stimuli accurately reflect the variability of the patterns of acquisition. The sessions depicted for SM 5 and SM 6 are not representative of the majority of their tests in which performance was near chance levels (see Table 1).

Acquisition with One Distinctive Stimulus Omitted

The previous tests indicated that when food was not available, the distinctive stimuli functioned to support rapid acquisition and that when food was available, removing all three distinctive stimuli delayed acquisition. These conditions did not reveal the specific reinforcing and/or discriminative functions of each distinctive stimulus for the three members of the chain when food was available, as during baseline sessions. To assess these functions, tests were conducted with one of the three distinctive stimuli omitted, but with the conditions preceding that link and following correct responses in the other links unaltered and with food available. Each kind of test condition was repeated three or four times with each subject. The portions of Figure 3 labelled "No 1st", "No 2nd", and "No 3rd" show the averaged results of these tests. The unfilled bars indicate the per cent correct without distinctive stimuli. The most apparent observation from these tests was that accuracy of all members of the chain deteriorated to some extent when the distinctive stimulus was removed for just one member of the chain. More central to the purposes of this experiment was the observation of changes in accuracy of the one member that was not followed by a distinctive

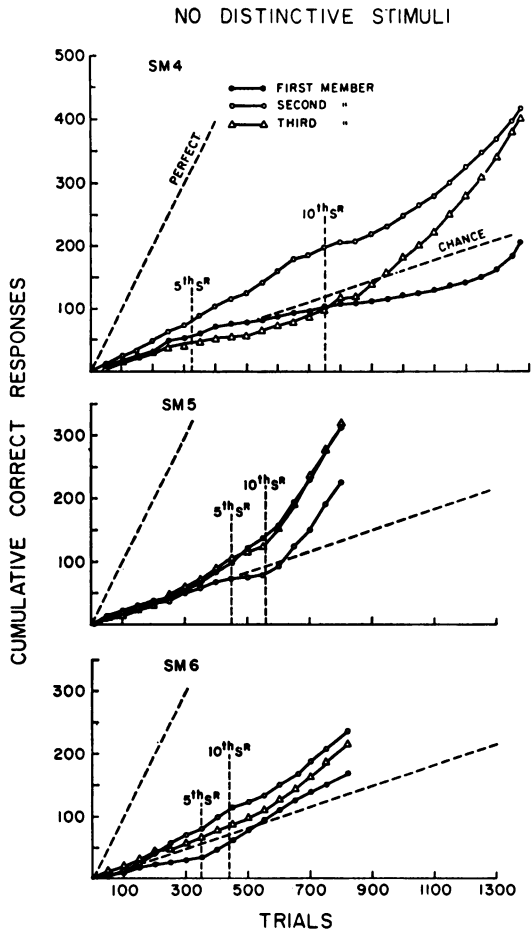


Fig. 5. Cumulative curves of correct responses in each link from representative completed sessions by the three subjects during the tests with no distinctive stimuli but with food. The curves were constructed like Figure 4, except that data were plotted across blocks of 50 trials. The fifth and tenth food reinforcers and slopes of perfect and chance performance are indicated.

stimulus. While there was considerable variability in accuracy between replications (see the lines extending from the bars in Figure 3), the ordering of accuracy among the three links of the chain for any single session was generally consistent with the averages shown in Figure 3. For all the subjects, acquisition of the member of the correct chain that was most severely retarded was the one that produced no distinctive stimulus. However, no retardation of acquisition was observed when the distinctive stimulus was omitted following the third member of the chain, the one immediately followed by food and food-cup illumination when a correct chain was emitted.

The pattern of acquisition within each session is depicted in Figure 6, which shows the accumulation of each of the three correct responses across trials. These curves of selected performances by SM 5 are generally representative of the results from all three subjects. In the session with the first distinctive stimulus omitted (top panel), the second and third members were acquired early in the session and were emitted with about 45% accuracy throughout the session. Acquisition of the first member was seriously retarded, showing an acceleration in occurrence only after about 1300 trials. In the session with the second distinctive stimulus omitted (middle panel), the first member was acquired first and occurred more reliably (68% correct) than the other members; the third member was acquired next and was slightly over 50% accurate throughout the session; acquisition of the second member was seriously retarded, showing acquisition only after 1350 trials. The bottom panel of Figure 6 shows performance when the third distinctive stimulus was omitted. The scales have been expanded to show the detail of this short session. No change in performance from baseline was apparent (compare to Figure 4 with all three distinctive stimuli and no food reinforcers); the accuracy of the first member of the chain was slightly superior to that of the other members. The performances with the first or second distinctive stimulus omitted shown in Figure 6 appear inferior to the performances shown in Figure 5 with no distinctive stimuli, because the latter were drawn from a nonrepresentative sample of tests in which acquisition occurred.

The rate of emitting chains was strongly related to the general level of accuracy. Table 1 shows that among the conditions with one distinctive stimulus omitted, overall accuracy (correct chains/total chains) and response rate (chains per minute) were lowest when the first distinctive stimulus was omitted, while accuracy and response rate were similar to baseline levels when the third distinctive stimulus was omitted. Since the rate of earning food pellets was a joint function of the accuracy and rate of responding, the rate of reinforcement was lowest when the distinctive stimulus was omitted after the first member of the chain, and was at baseline levels when omitted after the third member.

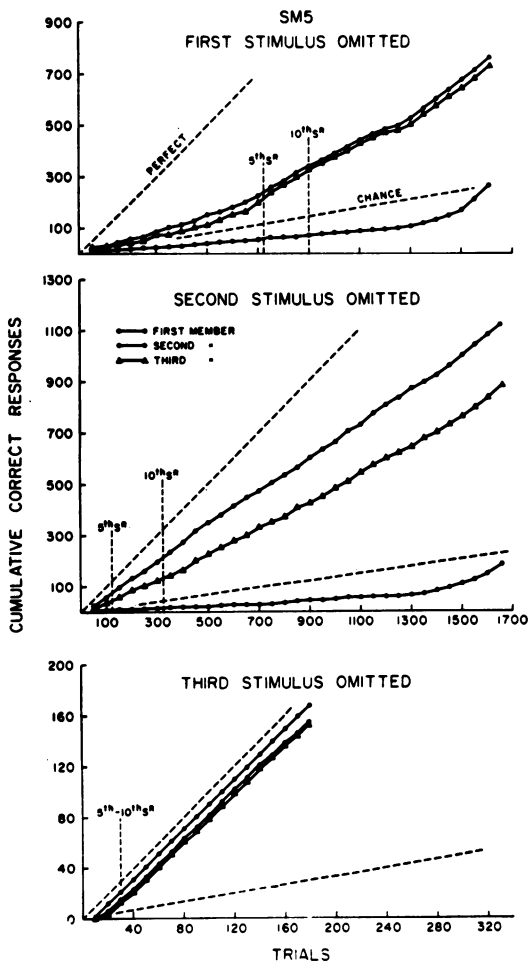


Fig. 6. Cumulative curves of correct responses in each link from representative sessions by SM 5 during the tests with one distinctive stimulus omitted. The curves were constructed like Figure 4. The fifth and tenth food reinforcers and slopes of perfect and chance performance are indicated.

Acquisition with a Distinctive Stimulus only after the First Member of the Chain

The selective effect of immediate differential consequences can be most dramatically seen if correct responses in only one link of the chain are followed by a distinctive stimulus. In the condition reported here, only the first member of the chain produced a distinctive stimulus. The first and second members followed the usual discriminative stimuli, but presentation of the distinctive stimulus after the first member improved accuracy, compared to the second member. Accuracy of the third member, which was occasionally followed by food presentation but did not follow the usual dis-

tinctive stimulus, was midway between the accuracy of the first and second members. Since accuracy of the third member was little affected by its distinctive stimulus, the results of this condition were similar to those of the "No 2nd" condition.

DISCUSSION

This study was designed to investigate the influence of differential consequences for correct responses on accuracy during acquisition. The various conditions indicated that a major function of these stimuli was to increase the probability of the members of the chain they followed, and for that reason they can be termed conditioned reinforcers. These same tests indicated that the distinctive stimuli strongly controlled the accuracy of subsequent members of the chain, and for that reason appeared to function as discriminative stimuli for correct responding in their presence. This dual function of the stimuli for correct responses is developed in the next two sections.

Reinforcement by Stimuli for Correct Responses

The test sessions conducted with no food reinforcers available to strengthen responses, but with the previously food-correlated stimuli presented after each correct response, demonstrated levels of accuracy and records of acquisition that were almost indistinguishable from sessions conducted with food available. Like previous demonstrations of acquisition of new behavior using only stimuli previously paired with primary reinforcement (e.g., Bersh, 1951; Clayton and Savin, 1960; Cowles, 1937; Crowder, Gill, Hodge, and Nash, 1959; Fox and King, 1961; Wolf, 1936) this demonstration showed that the stimuli functioned as reinforcers.

The increases in correct responses cannot be easily explained by other processes. Since a new response was acquired, an appeal to stimulus generalization or "elicitation" (see Kling and Schrier, 1971; Nevin, 1973) cannot account for the selection of the new response. As in a simultaneous discrimination, generalization among the alternatives would serve only to reduce accuracy. Nonspecific changes in the operant level or increases in general activity produced by occasional food-correlated stimuli would have had only a negligible effect on the

frequency of correct chains. Since 120 possible chains were available, the probability of an entire correct chain occurring by chance was less than 0.01; the random probability of 80 or more of them occurring in one session was virtually zero. The observation of acquisition in six independent test sessions without food reinforcement left no doubt that the presentation of distinctive stimuli after correct responses was conditioned reinforcement (see Kelleher and Gollub, 1962).

The high degree of control exerted by the stimuli for correct responses when no food was available might be traced to at least two aspects of the present procedure. First, many components of the acquisition process were not acquired anew during the tests. The conditions of each new session were the occasion for orienting, approach, and exploratory operation of the keys, which were part of the daily routine that eventually produced food. Although these conditions would not be expected to direct the subjects to acquire the specific new chain, it did predispose the subjects to emit many of the required responses early in the acquisition process, and increased the likelihood that the component responses would be emitted and so reinforced. This analysis bears directly on what has been called "learning to learn" (Blough and Lipsitt, 1971; Harlow, 1949; Meyer, 1960; Miles, 1957; Thompson, 1971).

Second, in contrast to previous attempts to condition a new response with conditioned reinforcers, the present procedure arranged for separate presentations of conditioned reinforcers for each component of the chain response, independent of the accuracy of the other components. This inherent "shaping" feature would be expected more rapidly to produce the terminal response, the correct chain. In this respect, the present procedure was similar to shaping a new response with tokens for primary reinforcement (Cowles, 1937).

To demonstrate that during baseline sessions the distinctive stimuli reinforced the correct response in the preceding link of the chain, tests were conducted in which only one or two members of the correct chain produced differential consequences. If the stimuli presented after correct responses reinforced those responses, then removing that reinforcer for one member of the chain should reduce the

probability of that member compared to the other members, and presenting the reinforcer for the first member only should produce greater accuracy for that member compared to the second and third. The results indicated that the distinctive stimuli functioned as reinforcers for the first and second members when food was available. However, omitting the distinctive stimulus following the third member did not disrupt accuracy.

The present results were in many ways comparable to the results of Boren (1969) and Boren and Devine (1968) in which a distinctive stimulus (timeout) presented after errors drastically reduced the probability of the responses they followed. One might say that their results were an example of differential punishment, while the present results were an example of differential reinforcement; however, the distinction between punishment and reinforcement in these experiments may be more apparent than real. In both cases, a different set of conditions followed correct and incorrect responses; it may be of little significance that in one case the distinctive feature was added after errors (punishment) and in the present case the distinctive feature was added after corrects (conditioned reinforcement). During discrimination training, a distinctive feature often exerts more control when provided on the positive discriminative stimulus (feature positive effect, see Jenkins and Sainsbury, 1969). More study is required to determine if a comparable asymmetry exists during acquisition between positive and negative stimulus consequences.

Stimulus Control by Stimuli for Correct Responses

The tests without food suggested that the distinctive stimuli reinforced all three members of the correct chain, since acquisition of the entire chain was rapid; yet, the tests with food but with no distinctive stimulus for the third member failed to decrease accuracy for the third member. This failure to observe evidence of conditioned reinforcement for the third member may not be surprising, given the immediacy of primary reinforcers for correct responses in the final position of the chain (see Egger and Miller, 1962). Accuracy of the third member was reduced, however, by changes in the usual distinctive stimuli *preceding* it, i.e., when the distinctive stimulus

did not follow the first or second member. For example, when the distinctive stimulus was omitted after the second member, accuracy for the third member was retarded more than accuracy for the first member that followed the usual discriminative stimulus (*i.e.*, the start of the trial with three yellow keys). Similar reductions in accuracy were observed for the second and third members when the distinctive stimulus did not follow the first member. These results indicate that the presence of distinctive stimuli controlled the accuracy of subsequent members of the chain, perhaps because during baseline sessions the presence of three such stimuli was the only occasion for food reinforcement of correct responses. Preventing the occurrence of the usual distinctive stimuli after the first or second member was equivalent to removing the discriminative stimulus for subsequent correct responses (see Mintz, Maurer, and Weinberg, 1966, for an analogous result with matching-to-sample).

Changes in Food Reinforcement Rate

One additional factor besides changes in conditioned reinforcement and stimulus control could have contributed to reduced performance during these tests. For example, when the distinctive stimulus did not follow the second member, accuracy of the *first* member was also lower than baseline levels. Since stimuli immediately preceding and following the first member of the chain were unaltered, neither changes in conditioned reinforcement nor stimulus control could be responsible. This result could be interpreted as a consequence of removing the delayed conditioned reinforcer for the second member, but was more probably related to the lower rate of food reinforcement in this condition compared to baseline. Both accuracy and rate of responding, the two factors controlling rate of reinforcement, were much lower than baseline when the second distinctive stimulus was omitted. For SM 5 and SM 6, some sessions were terminated before the entire daily food ration had been earned (after 18 hr). Since the rate of primary reinforcement was lower than baseline in several other tests as well, it is probable that this factor contributed somewhat to other decrements in accuracy and response rate. Note, however, that acquisition without food was nearly as rapid and accurate as during baseline and, further, that when the re-

quirements for food were not changed, large changes in accuracy were produced by changes in the available distinctive stimuli. Thus, the contingencies for food presentation probably played a minor role in generating the overall pattern of results.

The Pattern of Acquisition in Heterogeneous Chains

While the primary significance of these results was the demonstration of conditioned reinforcement during acquisition, a surprising feature of acquisition with conditioned reinforcement deserves mention. The extensive literature on chain and maze acquisition has suggested that the pattern of acquisition conforms more or less to a "goal gradient" (Arnold, 1947; Montpelier, 1933; *cf.* Sidman and Rosenberger, 1967; Spence and Shipley, 1934), *i.e.*, errors in performance are eliminated first near the reinforcer and last at positions remote from the reinforcer. Maze learning may be conceptualized as acquisition of a heterogeneous chain without presentation of *established* conditioned reinforcers for correct responses, a condition analogous to the condition in this experiment without any distinctive stimuli for correct responses. In a manner roughly analogous to the "goal gradient", the accuracy of the terminal response was slightly greater than the accuracy of the first response when acquisition was observed (see Figure 3). Similarly, in the condition with no conditioned reinforcers for the second and third member (1st only, Figure 3), accuracy of the third member was highest. However, during the baseline condition with established conditioned reinforcers for each correct response, no "goal gradient" was observed; in fact, the accuracy of the first response was on average greater than the accuracy of the last. Examining all conditions with distinctive stimuli for two or more chain members, without exception the average accuracy of the first of those members was the highest. Not only does this result suggest that observation of the goal gradient may require an absence of established conditioned reinforcers, it also suggests that the generalization that chains of behavior develop from the primary reinforcer backward is limited to cases in which established conditioned reinforcers are not available or not presented after the early members of the new chain.

REFERENCES

- Arnold, W. Simple reaction chains and their integration. I. Homogeneous chaining with terminal reinforcement. *Journal of Comparative and Physiological Psychology*, 1947, 40, 349-363.
- Bersh, P. The influence of two variables upon the establishment of a secondary reinforcer for operant responses. *Journal of Experimental Psychology*, 1951, 41, 62-73.
- Blough, D. and Lipsitt, L. The discriminative control of behavior, in J. W. Kling and L. A. Riggs (Eds), *Woodworth and Schlosberg's experimental psychology*. 3rd ed.; New York: Holt, Rinehart & Winston, 1971. Pp. 743-792.
- Boren, J. Repeated acquisition of new behavioral chains. *American Psychologist*, 1963, 17, 421. (abstract)
- Boren, J. Some variables affecting the superstitious chaining of responses. *Journal of the Experimental Analysis of Behavior*, 1969, 12, 959-969.
- Boren, J. and Devine, D. The repeated acquisition of behavioral chains. *Journal of the Experimental Analysis of Behavior*, 1968, 11, 651-660.
- Clayton, F. and Savin, H. Strength of a secondary reinforcer following continuous or variable ratio primary reinforcement. *Psychological Reports*, 1960, 6, 99-106.
- Cowles, J. Food-tokens as incentives for learning by chimpanzees. *Comparative Psychology Monographs*, 1937, 14, No. 5.
- Crowder, W., Gill, K., Jr., Hodge, C., and Nash, F., Jr. Secondary reinforcement or response facilitation? II. Response acquisition. *Journal of Psychology*, 1959, 48, 303-306.
- Egger, M. and Miller, N. Secondary reinforcement in rats as a function of information value and reliability of the stimulus. *Journal of Experimental Psychology*, 1962, 64, 97-104.
- Fox, R. and King, R. The effects of reinforcement scheduling on the strength of a secondary reinforcer. *Journal of Comparative and Physiological Psychology*, 1961, 54, 266-269.
- Harlow, H. The formation of learning sets. *Psychological Review*, 1949, 56, 51-65.
- Jenkins, H. and Sainsbury, R. The development of stimulus control through differential reinforcement. In N. J. Mackintosh and W. K. Honig (Eds), *Fundamental issues in associative learning*. Halifax: Dalhousie University Press, 1969. Pp. 133-161.
- Kelleher, R. and Gollub, L. A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 1962, 5, 543-597.
- Kling, J. and Schrier, A. Positive reinforcement. In J. W. Kling and L. A. Riggs (Eds), *Woodworth and Schlosberg's experimental psychology*. 3rd ed.; New York: Holt, Rinehart & Winston, 1971. Pp. 615-702.
- Meyer, D. The effects of differential probabilities of reinforcement on discrimination learning by monkeys. *Journal of Comparative and Physiological Psychology*, 1960, 53, 173-175.
- Miles, R. Learning-set formation in the squirrel monkey. *Journal of Comparative and Physiological Psychology*, 1957, 50, 356-357.
- Mintz, D., Mourer, D., and Weinberg, L. Stimulus control in fixed ratio matching-to-sample. *Journal of the Experimental Analysis of Behavior*, 1966, 9, 627-630.
- Montpellier, G., de. Note sur l'accélération dans les mouvements volontaire de la main. *Archives of Psychology*, Genève, 1937, 26, 181-197.
- Nevin, J. Conditioned reinforcement. In J. A. Nevin and G. S. Reynolds (Eds), *The study of behavior; learning, motivation, emotion and instinct*. Glenview, Illinois: Scott, Foresman and Company, 1973. Pp. 154-198.
- Sidman, M. and Rosenberger, P. Several methods for teaching serial position sequences to monkeys. *Journal of the Experimental Analysis of Behavior*, 1967, 10, 467-478.
- Skinner, B. *The behavior of organisms*. New York: Appleton-Century-Crofts, 1938.
- Spence, K. and Shipley, W. The factors determining the difficulty of blind alleys in maze learning by the white rat. *Journal of Comparative and Physiological Psychology*, 1934, 17, 423-436.
- Thompson, D. Transition to a steady state of repeated acquisition. *Psychonomic Science*, 1971, 24, 236-238.
- Thompson, D. Repeated acquisition of response sequences: stimulus control and drugs. *Journal of the Experimental Analysis of Behavior*, 1975, 23, 429-436.
- Wolfe, J. Effectiveness of token rewards for chimpanzees. *Comparative Psychology Monographs*, 1936, 12, No. 60.

Received 12 February 1976.

(Final Acceptance 1 October 1976.)